Seminars

Date	Speaker	Seminar
9 March 2020 Room M/0.40	Almut Veraart (Imperial College London)	Volatility estimation in time and space The concept of (stochastic) volatility/intermittency is of central importance in many fields of science. In this talk I am going to discuss how stochastic volatility can be introduced in a stochastic model and which properties of the stochastic model have an influence on the methods available for volatility estimation. I will showcase some recent results on how stochastic volatility can be estimated in multivariate non-semimartingale settings and show some first results in extending the classical stochastic volatility concept to spatial/spatio-temporal settings. The results presented in this talk are based on collaborations with Ole E. Barndorff-Nielsen, Fred Espen Benth, Andrea Granelli, Michele Nguyen, Riccardo Passaggeri.
2 March 2020	Ioannis Kosmidis (Warwick University)	Improved estimation of models for ordinal responses For the estimation of cumulative link models and adjacent category models for ordinal data, we derive adjustments to the likelihood score functions, whose solution ensures an estimator with smaller asymptotic bias than the maximum likelihood estimator typically has. The form of the adjustments suggests a parameter-dependent adjustment of the multinomial counts, which in turn suggests the solution of the adjusted score equations through iterated maximum likelihood fits on adjusted counts, greatly facilitating implementation. Like the maximum likelihood estimator, the reduced-bias estimator is found to respect the key invariance properties that make cumulative link models a good choice for the analysis of categorical data. Its additional finiteness and optimal frequentist properties, along with the adequate behaviour of related asymptotic inferential procedures, make the reduced-bias estimator attractive as a default choice for practical applications.

		We will also discuss the improved estimation of the adjacent category model, which is another popular model for ordinal data, and how this can be achieved using a modification of the so-called "Poisson trick".
13 February 2020 Time:14:10 to 15:10 Room M/2.20	Tatiana Benaglia (University of Campinas)	Bayesian Mixture Models for longitudinal data on cognition loss in elderly people A regression mixture model to handle elderly's cognitive ability up to their death is presented. Cognition is measured across time with standard questionnaires from geriatrics which involve, amongst others, memory, language and reasoning issues. The output of such questionnaires is recorded with a countable and finite score. Models for Binomial response variables are discussed here. The mixture specification rises to discriminate two prevalent behaviours in the data: one group of elderly people presents cognition decline at constant rate; whilst the other experiences a spontaneous accelerated decline at some time. The latter aspect is dealt with random change points nonlinear predictors. In addition, logit and complementary log-log link functions were used to model the mixture allocation with predictor variables. The study's goal is to quantify associations amidst cognition loss and the diagnostics of dementias like Alzheimer's disease, besides sociodemographic factors. The proposed model is evaluated in the database provided by the Rush University - Chicago, United States, through the Rush Memory and Aging Project from 1997 to 2016. The talk is based on joint work with Eric Krishna, Hildete P. Pinheiro (Campinas) and Graciela Muniz-Terrera (Edinburgh).
10 February 2020	Xin Liu (University of Bath)	Diversification in Lottery-Like Features and Portfolio Pricing Discounts I study the asset pricing implications of cumulative prospect theory on portfolio discounts. I extend Barberis and Huang (2008) and show that a portfolio consisting of lottery-like stocks should trade at a discount due to diversification. This discount can be partially mitigated if lottery-like stocks tend to produce extreme payoffs at the same time. I utilize three empirical settings to support this theoretical prediction: the closed-end fund puzzle, the announcement returns of mergers and acquisitions, and conglomerate discounts. My findings support cumulative prospect theory from an alternative perspective and provide a novel and unifying explanation for three seemingly unrelated phenomena.
27 January 2020	Dino Sejdinovic (University of Oxford)	Noise Contrastive Meta-Learning for Conditional Density Estimation using Kernel Mean Embeddings Current meta-learning approaches focus on learning functional representations of relationships between variables, i.e. estimating

		conditional expectations in regression. In many applications, however, the conditional distributions cannot be meaningfully summarized solely by expectation (due to e.g. multimodality). We introduce a novel technique for meta-learning conditional densities, which combines neural representation and noise contrastive estimation together with well-established literature in conditional mean embeddings into reproducing kernel Hilbert spaces. The method shows significant improvements over standard density estimation methods on synthetic and real-world data, by leveraging shared representations across multiple conditional density estimation tasks.
17 December 2019	Eliana Christou (UNC Charlotte, North Carolina)	Central Quantile Subspace Quantile regression (QR) is becoming increasingly popular due to its relevance in many scientific investigations. There is a great amount of work about linear and nonlinear QR models. Specifically, nonparametric estimation of the conditional quantiles received particular attention, due to its model flexibility. However, nonparametric QR techniques are limited in the number of covariates. Dimension reduction offers a solution to this problem by considering low-dimensional smoothing without specifying any parametric or nonparametric regression relation. Existing dimension reduction techniques focus on the entire conditional distribution. We, on the other hand, turn our attention to dimension reduction techniques for conditional quantiles and introduce a new method for reducing the dimension of the predictor X. The novelty of this paper is threefold. We start by considering a single index quantile regression model, which assumes that the conditional quantile depends on X through a single linear combination of the predictors, then extend to a multi index quantile regression model, and finally, generalize the proposed methodology to any statistical functional of the conditional distribution. The performance of the methodology is demonstrated through simulation examples and real data applications. Our results suggest that this method has a good finite sample performance and often outperforms existing methods. Please note this talk takes place at 11:10 in room M/2.06.
9 December 2019	Ruth Misener (Imperial College)	Scoring positive semidefinite cutting planes for quadratic optimization via trained neural networks Semidefinite programming relaxations complement polyhedral relaxations for quadratic optimization, but global optimization solvers built on polyhedral relaxations cannot fully exploit this advantage. We develop linear outer-approximations of semidefinite constraints that can be effectively integrated into global solvers for nonconvex quadratic optimization. The difference from previous work is that our proposed cuts are (i) sparser with respect to the number of

		nonzeros in the row and (ii) explicitly selected to improve the objective. A neural network estimator is key to our cut selection strategy: ranking each cut based on objective improvement involves solving a semidefinite optimization problem, but this is an expensive proposition at each Branch&Cut node. The neural network estimator, trained a priori of any instance to solve, takes the most time consuming computation offline by predicting the objective improvement for any cut. This is joint work with Radu Baltean-Lugojan, Pierre Bonami, and Andrea Tramontani.
2 December 2019	Edilson Fernandes De Arruda (Rio de Janeiro / Cardiff)	Solving Markov Processes by Time Aggregation: Theory and Applications Markov decision processes are a natural way to model sequential decisions under uncertainty and find applications in many fields, such as healthcare, renewable energy and supply chains. Unfortunately, complex problems give rise to very large state spaces (curse of dimensionality), rendering classical algorithms intractable. In this talk, I will discuss some algorithms that make use of time aggregation (embedding) to tackle the curse of dimensionality and seek optimal or sub-optimal solutions to complex systems in reasonable computational time. I will present some real-world applications to illustrate both the approach and the flexibility of Markov decision processes as a modelling tool.
25 November 2019	Theo Economou (University of Exeter)	An Advanced Hidden Markov Model for Hourly Rainfall Time Series For hydrological applications, such as urban flood modelling, it is often important to be able to simulate sub-daily rainfall time series from stochastic models. However, modelling rainfall at this resolution poses several challenges, including a complex temporal structure including long dry periods, seasonal variation in both the occurrence and intensity of rainfall, and extreme values. We illustrate how the hidden Markov framework can be adapted to construct a compelling model for sub-daily rainfall, which is capable of capturing all of these important characteristics well. These adaptations include clone states and non-stationarity in both the transition matrix and conditional models. Set in the Bayesian framework, a rich quantification of both parametric and predictive uncertainty is available, and thorough model checking is made possible through posterior predictive analyses. Results from the model are interpretable, allowing for meaningful examination of seasonal variation and medium to long term trends in rainfall occurrence and intensity. To demonstrate the effectiveness of our approach, both in terms of model fit and interpretability, we apply the model to an 8-year long time series of hourly observations.

18 November 2019	Jack Noonan (Cardiff University)	First passage time for Slepian process with linear barrier In 1971, L.A. Shepp found explicit formulas for the first passage probability Pr(S(t)<a for all t in [0,T] | S(0)=x), for all T>0, where S(t)is a Gaussian process with mean 0 and covariance E S(t)S(t')=max{0,1-|t-t'|}. In a recent paper, we extended Shepp's results to the more general case of piecewise-linear barriers; previously, explicit formulas for even Pr(S(t)<a+bt for all t in [0,T]) were known only for the cases b=0 (constant barrier) or T <= 1 (short interval). In this talk, we outline applications to a change point detection problem; detecting temporary drift change of Brownian. After discussing Average Run Length (ARL) approximations, we formulatevery accurate approximations for the power of the test. We also investigate the performance of the test when the change in drift is permanent and compare performance to the known optimal CUSUM and Shiryaev-Roberts procedures.
11 November 2019	Enrica Pirozzi (University of Naples)	On a Fractional Ornstein-Uhlenbeck Process and its applications The seminar is centred on a fractional Ornstein-Uhlenbeck process that is solution of a linear stochastic differential equation, driven by a fractional Brownian motion; it is also characterised by a stochastic forcing term in the drift. For such a process, mean and covariance functions will be specified, concentrating on their asymptotic behaviour. A sort of short- or long-range dependence, under specified hypotheses on the covariance of the forcing process, will be shown. Applications of this process in neuronal modelling are discussed, providing an example of a stochastic forcing term as a linear combination of Heaviside functions with random center. Simulation algorithms for the sample path of this process are also given.
4 November 2019	Emma Aspland (Cardiff University)	Lung Cancer Clinical Pathway Modelling Clinical pathways are an effective and efficient approach in standardising the progression of treatment to support patient care and facilitate clinical decision making. Our review of the related literature highlighted a need to better integrate data engineering and OR techniques with expert/domain knowledge to assist with clinical pathway discovery and formation. Consequently, we have produced a decision support tool that facilitates expert interaction with data mining, through the application of clustering. This has involved the development of a new distance metric, modified from the Needleman-Wunsch algorithm, that considers weightings and groupings of activities as specified by an expert user. The resulting set of pathways are then automatically translated into the basis of a

		discrete event simulation to model patient flows through the captured clinical pathways.
4 November 2019	Clement Twumasi (Cardiff University)	Comparative modelling of parasite population dynamics of two Gyrodactylus species Understanding fully host-parasite systems is challenging if employing just experimental approaches, whereas mathematical models can help uncover in-depth knowledge of the infection dynamics. The current study compares the infection dynamics of two parasite species (Gyrodactylus turnbulli and Gyrodactylus bullatarudis) across three host populations (ornamental, Lower Aripo and Upper Aripo fish), by developing a Continuous-time Markov Chain (CTMC) model. The model simulates the movement of parasites for two age groups over the external surfaces (eight body regions) of a fish over a 17-day infection period with population carrying capacity (dependant on host size and area of body regions). The model was parameterised by the birth, death and movement rates of young and older parasites, in the presence or absence of host's immune response. Host death was assumed to occur at a rate proportional to the total number of parasites on the fish. The CTMC simulation model was fitted using a novel Weighted-iterative Approximate Bayesian Computation (ABC). The findings from this study would help policy makers and biologists to better understand the Gyrodactylus-fish system using mathematical models and inform management decisions for the control of gyrodactylid infections.
21 October 2019	Tri-Dung Nguyen (University of Southampton)	Game of Banks – Keeping Free ATMs Alive? The LINK ATM network is a fundamental part of the UK's payments infrastructure - with nearly 62,000 ATMs - and cash machines are by far the most popular channel for cash withdrawal in the UK, used by millions of consumers every week. The record high daily withdrawal in 2019 was 10.7 million ATM transactions (29 March) and with over half a billion pounds paid out by ATMs (28 June). The UK's cash machine network is special in that most of them are currently free of charge. Underlying this key feature is the arrangement among the banks and cash machine operators to settle the fees among themselves instead of putting the burden on the consumers' shoulders. The ATM network in the UK has recently, however, been experiencing many issues as some members are not happy with the mechanism for interchange fee sharing. In this talk, we show how Game Theory, especially how to combine mathematical models developed by John

		Nash and Lloyd Shapley, two Nobel laureates in Economics, to resolve the current ATM crisis. We present a novel `coopetition' game theoretic model for banks to optimally invest in the ATM network and to share the cost. This coopetition game includes both a cooperative game theory framework as the mechanism for interchange fee sharing and a non-cooperative counterpart to model the fact that banks also wish to maximise their utilities. We show that the current mechanism for sharing is unstable, which explains why some members are threatening to leave. We also show that, under some settings, the Shapley allocation belongs to the core and hence it is not only fair to all members but also leads to a stable ATM network. We prove the existence of a pure Nash equilibrium, which can be computed efficiently. In addition, we show that the Shapley value allocation dominates the current mechanism in terms of social welfare. Finally, we provide numerical analysis and managerial insights through a case study using real data on the complete UK ATM network.
14 October 2019	Ruth King (University of Edinburgh)	Challenges of quantity versus complexity for ecological data Capture-recapture data are often collected on animal populations to obtain insight into the given species and/or ecosystem. Long-term datasets combined with new technology for observing individuals are producing larger capture-recapture datasets – for example, repeated observations on >10,000 individuals are becoming increasingly common. Simultaneously, increasingly complex models are being developed to more accurately represent the underlying biological processes which permit a more intricate understanding of the system. However, fitting these more complex models to large datasets can become computationally very expensive. We propose a two step Bayesian approach: (i) fit the given capture-recapture model to a smaller subsample of the data; and then (ii) “correct” the posterior obtained so that it is (approximately) from the posterior distribution of the complete sample. For a feasibility study we apply this two-step approach to data from a colony of guillemots where there are approximately 30,000 individuals observed within the capture-recapture study and investigate the performance of the algorithm.
7 October 2019	George Loho (LSE)	Tropical volume by tropical ehrhart polynomials We are motivated by the problem of intrinsically defining a volume concept for tropical polytopes. In Euclidean space, the number of integral points contained in a large dilate of a given polytope approximates its volume. Even more, by Ehrhart's theorem, the

		function that counts integral points in dilates of an integral polytope is a polynomial in the dilation factor, whose leading coefficient equals the volume of the polytope. After introducing a suitable and natural concept of a tropical lattice, we aim to establish an analogous result for tropical lattice polytopes based on their covector decomposition. We compare the arising volume notion with existing ones and demonstrate basic properties. This is based on joint work with Matthias Schymura (EPFL).
30 September 2019	Rajen Shah (University of Cambridge)	RSVP-graphs: Fast High-dimensional Covariance Matrix Estimation Under Latent Confounding We consider the problem of estimating a high-dimensional p × p covariance matrix S, given n observations of confounded data with covariance S + GG^T , where G is an unknown p × q matrix of latent factor loadings. We propose a simple and scalable estimator based on the projection on to the right singular vectors of the observed data matrix, which we call RSVP. Our theoretical analysis of this method reveals that in contrast to approaches based on removal of principal components, RSVP is able to cope well with settings where the smallest eigenvalue of G^T G is relatively close to the largest eigenvalue of S, as well as when eigenvalues of G^T G are diverging fast. RSVP does not require knowledge or estimation of the number of latent factors q, but only recovers S up to an unknown positive scale factor. We argue this suffices in many applications, for example if an estimate of the correlation matrix is desired. We also show that by using subsampling, we can further improve the performance of the method. We demonstrate the favourable performance of RSVP through simulation experiments and an analysis of gene expression datasets collated by the GTEX consortium.